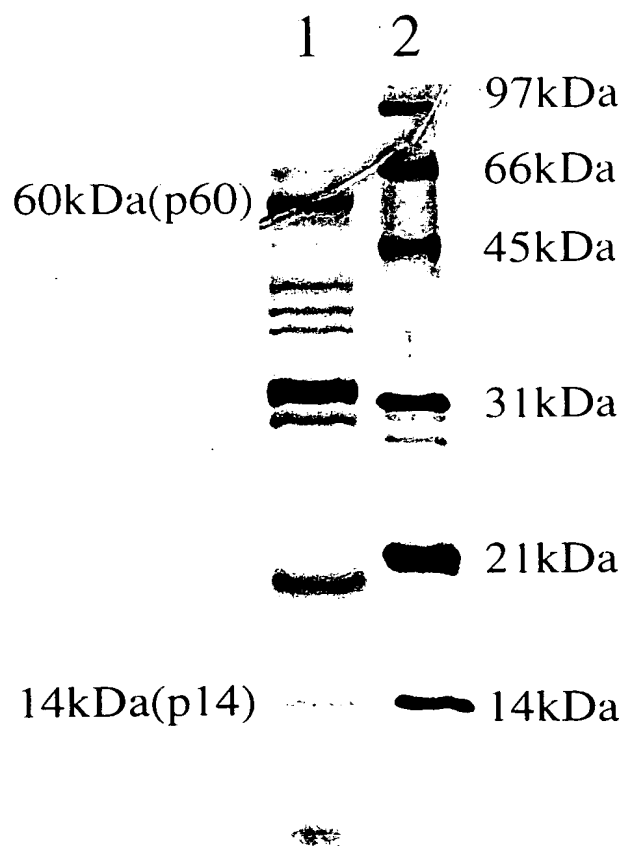




FIG. 1

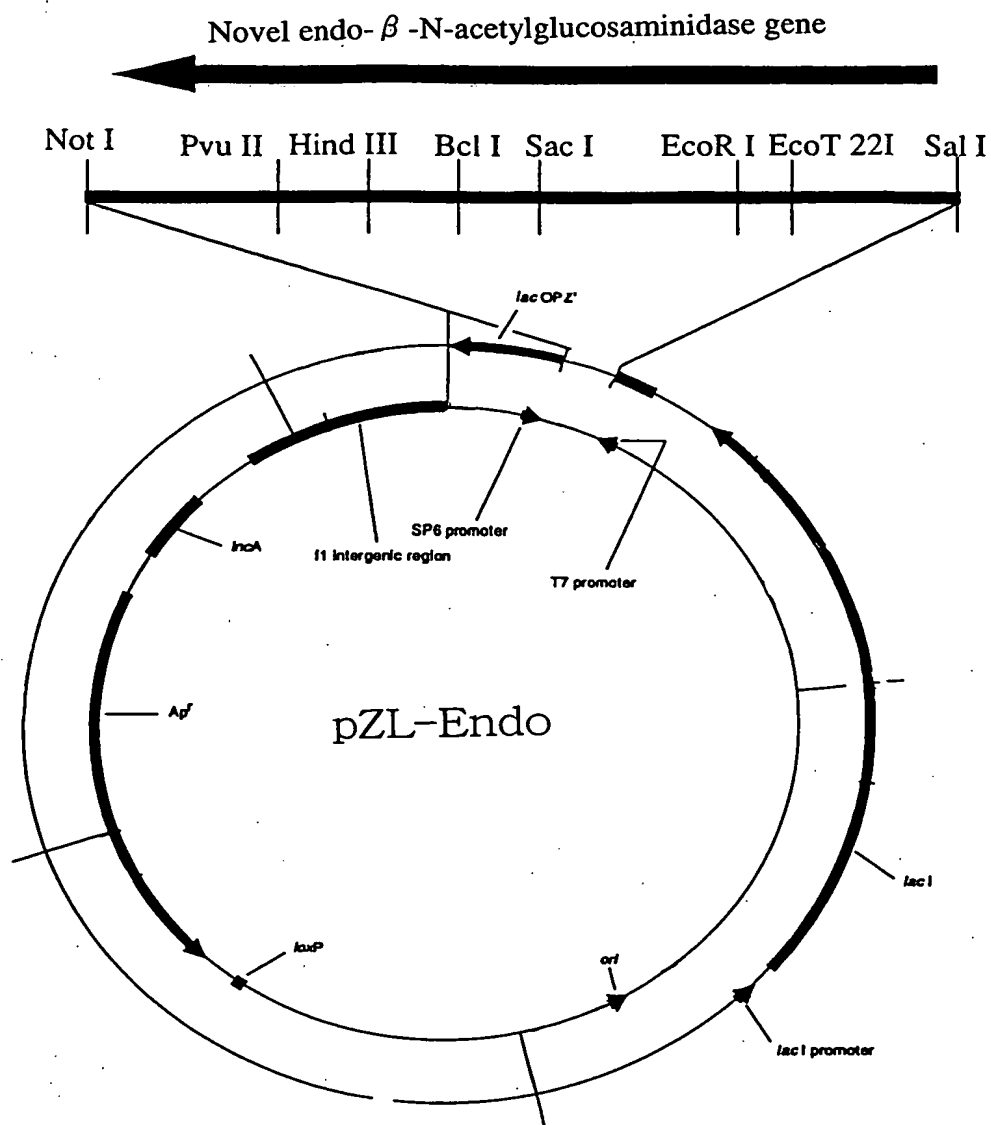


Purification Result for Endo- $\beta$ -N-acetylglucosaminidase (15-25% gradient SDS-PAGE)

Lane 1: Purified endo- $\beta$ -N-acetylglucosaminidase from *Mucor hiemalis*

Lane 2: Molecular weight markers

FIG. 2



Restriction enzyme map for pZL-Endo including the full-length novel endo- $\beta$ -N-acetylglucosaminidase gene.

# FIG. 3



10	20	30	40	50	60
GTCGACCCAC	GGGTCCGCG	ACGCGTGGGC	GGACGCGTGG	GCGGACGCGT	GGGTTTTATT
70	80	90	100	110	120
TTACATAAAT	ATGCCTTCAC	TTCAATTGCA	ACCTGATGAC	AAACTAGCAC	CTGTTTCTTT
130	140	150	160	170	180
TGCACTTAAG	TCTATGAATG	AGTTGAGGGA	CTGGACGCCA	GACGAAAAGA	TAAAGTTTAA
190	200	210	220	230	240
CGTTTCAAGC	GTGGCACTAC	AGCCTCGTGT	GAAAAACGCC	CTGAAACCTC	AATTATTTGT
250	260	270	280	290	300
AACTCATGAT	ATGGCAGGAG	GATATAAAGA	AGATAAAAAT	ATTCAAGGAA	ACAATTATAA
310	320	330	340	350	360
AGACATTTAT	AACATTCAAT	ATTGGCAITTT	AGCTGATACT	TTTGTATATT	TCTCTCATGA
370	380	390	400	410	420
GCGAGTTAGC	ATTCTCTCAG	TCAATTGGAC	AAATGCTTGT	CATAGAAATG	GTGTAAAGTG
430	440	450	460	470	480
TTTAGGTACT	TTTTTAGTAG	AAGGAAATAA	CCAAATGCAT	GAAATGGAAG	CCTTGCTTCA
490	500	510	520	530	540
CGGTCCACCT	TACTTTAATA	ACACTGACGA	CCCTATGAGA	TTATGGAGTC	CGTATTATGC
550	560	570	580	590	600
AGACCAATTA	GTTGCTATTG	CTAAACACTA	TGGTTTTGAT	GGCTGGTTGT	TCAATATTGA
610	620	630	640	650	660
ATGCGAATTC	TTTCCTTTTC	CTACAAATCC	AAAATTCAAA	GCTGAAGAGT	TGGCAAAGTT
670	680	690	700	710	720
TCTACACTAT	TTTAAGGAAA	AATTGCATAA	CGAAATACCT	GGATCTCAAC	TCATTGGTA
730	740	750	760	770	780
CGACAGCATG	ACAAATGAAG	GAGAAATCCA	CTGGCAGAAC	CAGCTCACAT	GGAAAAATGA
790	800	810	820	830	840
GTTATTTTTT	AAAAACACGG	ATGGTATTTT	TTTGAATTAT	TGGTGGAAAA	AAGAATACCC
850	860	870	880	890	900
TGAAATGGCG	CGTAGAGTAG	CTGAAGGAAT	AGGTAGATCA	GGTTTAGAAG	TTTATTTTGG
910	920	930	940	950	960
TACAGATGTA	TGGGAAGGC	ATACTTATGG	TGGCGGTGGT	TTCAAATCAT	ATAAGGGTGT
970	980	990	1000	1010	1020
AAAAACTGCC	TACTCTGCAA	TGACATCTTC	TGCATTATTT	GGTATGGCAT	GGACATACGA
1030	1040	1050	1060	1070	1080
GCATTTTCGAA	AAGTCTGAAT	TTGAAAAGAT	GGATCGTTTG	TTTTGGTGTG	GTGGTAAATA
1090	1100	1110	1120	1130	1140
CTCTGACTAT	CCTCCCCCAC	CTCCTAAAAA	CCCAGATGAC	GAAAAAGAAG	TAGAAAGCGA
1150	1160	1170	1180	1190	1200
TGATAGTGAA	GATGAGCTCA	TGTACGGACA	CAAGAAAGGT	ATTGCTGACA	CGGTAGAATC
1210	1220	1230	1240	1250	1260
TATTCCTGTA	CCAGGAACAG	ATTGGTTTGT	TACCAATTTT	GATAGGGGGT	TTGGAAATAG
1270	1280	1290	1300	1310	1320
GTTTTATTAT	AGAGGAAAGA	GATTACTTTC	TCAGCCTTGG	TCCCATTTTAT	CGCATCAAGC
1330	1340	1350	1360	1370	1380
TATTCTCCCC	AATAAAGCT	ATCGAAATCC	AGAGATTTAT	CCCACTGATC	AAAACATTAA

Entire nucleotide sequence of the fragment inserted into the Sal I-Not I sites of pZL-Endo including the full-length novel endo- $\beta$ -N-acetylglucosaminidase gene.

# FIG. 4



1390	1400	1410	1420	1430	1440
AATCACTAGT	TCTCTCGATT	GCGATCATGG	AGCTTTTCTT	GGTGGAACTT	CGCTTATTAT
1450	1460	1470	1480	1490	1500
CAAAGGCCAA	CGTTTCAATC	ATAGAGAATC	GCATGATGTT	GAAACTGAAA	TTAGTATACC
1510	1520	1530	1540	1550	1560
TCTGTATAAG	CTTTTCATTAG	ATGCTAGTAA	AGGATGCTCA	TTGCGTTATA	TTTATAGAAC
1570	1580	1590	1600	1610	1620
TTTGTGTATG	AAAGATGTAA	AGTTGACAGT	AGCATGTCAC	TTTTTCGTTAA	AAACAAACGA
1630	1640	1650	1660	1670	1680
CTCAGTTAAT	TTCTTCAAGG	TATGGCAGCC	AGATGAAAAT	TTCTCTTTTG	AATATGATGA
1690	1700	1710	1720	1730	1740
TGGAATGAGA	GCCACTGTTA	CAACTGAAAA	TTCTACCGAA	AGCAGATGCT	TTTTATTACG
1750	1760	1770	1780	1790	1800
TACAACAGAA	GAAGATACAG	GAGAAAATGA	TTGGATAACA	AAAACATTATA	ATGTGCTGTC
1810	1820	1830	1840	1850	1860
TGTTCCAGAA	GGAAGTCAAT	TATACATTAC	AAGACTTGAA	GTGAGCGTAG	TATTAGATAC
1870	1880	1890	1900	1910	1920
AGCTGGTTTA	GTAGGICTTG	TTAATCAAGT	TATTGCTTGC	TTGGGATATA	TTAGCATCAT
1930	1940	1950	1960	1970	1980
ACCAACTATA	AATTCCTGAA	TAAAAACAGA	TTCATCACGC	ATTATTTCAGG	ATCTCTTTTG
1990	2000	2010	2020	2030	2040
GAAAGATCAG	AAATATACCA	AAATCGGAAA	AGAAAGTTTA	GACGACATAG	CTCAAGAAGA
2050	2060	2070	2080	2090	2100
AGTTCATAGA	TATTATGGAA	CATTGAACTG	GGAAAAACACA	GCAAATGTAG	TAAACGCTTG
2110	2120	2130	2140	2150	2160
GGAGGAAATA	GATTACTACA	ACGTTTTTTTA	CAAAGAAAGT	GACGACTCTG	CAACTCGCAT
2170	2180	2190	2200	2210	2220
CTTTTTTAGGA	ACAGCATTCT	GTAATCAATT	TCGTGTATCT	GGTTTAGATA	TTATTTTATC
2230	2240	2250	2260	2270	2280
TAAGCTACCA	AAGATAGTTA	TTGAAGCTGT	TAACAAAGAA	GGATACATCT	CTTCAAGTGG
2290	2300	2310	2320	2330	2340
TAGCATAGAT	TTGTCATTAA	ACTAGGACTT	GAAATAAAAT	ATTATGATAA	AGAAAAAAA
2350	2360	2370	2380	2390	2400
AAAAAAAAAA	AAAAAAAAAG	GGCGGCCGC.	.....	.....	.....

Entire nucleotide sequence of the fragment inserted into the Sal I-Not I sites of pZL-Endo including the full-length novel endo- $\beta$ -N-acetylglucosaminidase gene. (Continued)

# FIG. 5

5'	ATG	CCT	TCA	CTT	CAA	TTG	CAA	CCT	GAT	GAC	AAA	CTA	GCA	CCT	GTT	TCT	TTT	GCA	54
	M	P	S	L	Q	L	Q	P	D	D	K	L	A	P	V	S	F	A	
	CTT	AAG	TCT	ATG	AAT	GAG	TTG	AGG	GAC	TGG	ACG	CCA	GAC	GAA	AAG	ATA	AAG	TTT	108
	L	K	S	M	N	E	L	R	D	W	T	P	D	E	K	I	K	F	
	AAC	GTT	TCA	AGC	GTG	GCA	CTA	CAG	CCT	CGT	GTG	AAA	AAC	GCC	CTG	AAA	CCT	CAA	162
	N	V	S	S	V	A	L	Q	P	R	V	K	N	A	L	K	P	Q	
	TTA	TTG	TTA	ACT	CAT	GAT	ATG	GCA	GGA	GGA	TAT	AAA	GAA	GAT	AAA	AAT	ATT	CAA	216
	L	L	L	T	H	D	M	A	G	G	Y	K	E	D	K	N	I	Q	
	GGA	AAC	AAT	TAT	AAA	GAC	ATT	TAT	AAC	ATT	CAA	TAT	TGG	CAT	TTA	GCT	GAT	ACT	270
	G	N	N	Y	K	D	I	Y	N	I	Q	Y	W	H	L	A	D	T	
	TTT	GTA	TAT	TTC	TCT	CAT	GAG	CGA	GTT	AGC	ATT	CCT	CCA	GTC	AAT	TGG	ACA	AAT	324
	F	V	Y	F	S	H	E	R	V	S	I	P	P	V	N	W	T	N	
	GCT	TGT	CAT	AGA	AAT	GGT	GTA	AAG	TGT	TTA	GGT	ACT	TTT	TTA	GTA	GAA	GGA	AAT	378
	A	C	H	R	N	G	V	K	C	L	G	T	F	L	V	E	G	N	
	AAC	CAA	ATG	CAT	GAA	ATG	GAA	GCC	TTG	CTT	CAC	414	CCA	CCT	TTA	CTT	AAT	AAC	432
	N	Q	M	H	E	M	E	A	L	L	H	G	P	P	L	L	N	N	
	ACT	GAC	GAC	CCT	ATG	AGA	TTA	TGG	AGT	CCG	TAT	TAT	GCA	GAC	CAA	TTA	GTT	GCT	486
	T	D	D	P	M	R	L	W	S	P	Y	Y	A	D	Q	L	V	A	
	ATT	GCT	AAA	CAC	TAT	GGT	TTT	GAT	GGC	TGG	TTG	TTT	AAT	ATT	GAA	TGC	GAA	TTC	540
	I	A	K	H	Y	G	F	D	G	W	L	F	N	I	E	C	E	F	
	TTT	CCT	TTT	CCT	ACA	AAT	CCA	AAA	TTT	AAA	GCT	GAA	GAG	TTG	GCA	AAG	TTT	CTA	594
	F	P	F	P	T	N	P	K	F	K	A	E	E	L	A	K	F	L	
	CAC	TAT	TTT	AAG	GAA	AAA	TTG	CAT	AAC	GAA	ATA	CCT	GGA	TCT	CAA	CTC	ATT	TGG	648
	H	Y	F	K	E	K	L	H	N	E	I	P	G	S	Q	L	I	W	
	TAC	GAC	AGC	ATG	ACA	AAT	GAA	GGA	GAA	ATC	CAC	684	TGG	CAG	AAC	693	CTC	ACA	702
	Y	D	S	M	T	N	E	G	E	I	H	W	Q	N	Q	L	T	W	

Amino acid sequence deduced from the novel Endo- $\beta$ -N-acetylglucosaminidase gene, and the nucleotide sequence of the DNA encoding this amino acid sequence.

# FIG. 6



711	720	729	738	747	756
AAA AAT GAG TTA TTT	TTT AAA AAC ACG GAT GGT	ATT TTT TTG AAT TAT TGG	TGG		
K N E L F F	K N T D G I F L N Y W W				
765	774	783	792	801	810
AAA AAA GAA TAC CCT	GAA ATG GCG CGT AGA GTA GCT	GAA GGA ATA GGT AGA TCA			
K K E Y P E M A R R V A E G I G R S					
819	828	837	846	855	864
GGT TTA GAA GTT TAT	TTT GGT ACA GAT GTA TGG	GGA AGG CAT ACT TAT GGT	GGC		
G L E V Y F G T D V W G R H T Y G G					
873	882	891	900	909	918
GGT GGT TTC AAA TCA TAT	AAG GGT GTA AAA ACT GCC	TAC TCT GCA ATG ACA TCT			
G G F K S Y K G V K T A Y S A M T S					
927	936	945	954	963	972
TCT GCA TTA TTT GGT	ATG GCA TGG ACA TAC GAG	CAT TTC GAA AAG TCT GAA	TTT		
S A L F G M A W T Y E H F E K S E F					
981	990	999	1008	1017	1026
GAA AAG ATG GAT CGT TTG	TTT TGG TGT GGT GGT	AAA TAC TCT GAC TAT CCT	CCC		
E K M D R L F W C G G K Y S D Y P P					
1035	1044	1053	1062	1071	1080
CCA CCT CCT AAA AAC CCA	GAT GAC GAA AAA GAA GTA	GAA AGC GAT GAT AGT	GAA		
P P P K N P D D E K E V E S D D S E					
1089	1098	1107	1116	1125	1134
GAT GAG CTC ATG TAC GGA	CAC AAG AAA GGT ATT GCT	GAC ACG GTA GAA TCT	ATT		
D E L M Y G H K K G I A D T V E S I					
1143	1152	1161	1170	1179	1188
CCT GTA CCA GGA ACA GAT	TGG TTT GTT ACC AAT TTT	GAT AGG GGG TTT GGA	AAT		
P V P G T D W F V T N F D R G F G N					
1197	1206	1215	1224	1233	1242
AGG TTT TAT TAT AGA GGA	AAG AGA TTA CTT TCT	CAG CCT TGG TCC CAT	TTA TCG		
R F Y Y R G K R L L S Q P W S H L S					
1251	1260	1269	1278	1287	1296
CAT CAA GCT ATT CTC CCC	AAT AAA AGC TAT CGA	AAT CCA GAG ATT TAT	CCC ACT		
H Q A I L P N K S Y R N P E I Y P T					
1305	1314	1323	1332	1341	1350
GAT CAA AAC ATT AAA ATC	ACT AGT TCT CTC GAT	TGC GAT CAT GGA GCT	TTT CTT		
D Q N I K I T S S L D C D H G A F L					
1359	1368	1377	1386	1395	1404
GGT GGA ACC TCG CTT ATT	ATC AAA GGC CAA CGT	TTC AAT CAT AGA GAA	TCG CAT		
G G T S L I I K G Q R F N H R E S H					
1413	1422	1431	1440	1449	1458
GAT GTT GAA ACT GAA ATT	AGT ATA CCT CTG TAT	AAG CTT TCA TTA GAT	GCT AGT		
D V E T E I S I P L Y K L S L D A S					

Amino acid sequence deduced from the novel Endo- $\beta$ -N-acetylglucosaminidase gene, and the nucleotide sequence of the DNA encoding this amino acid sequence. (Continued)

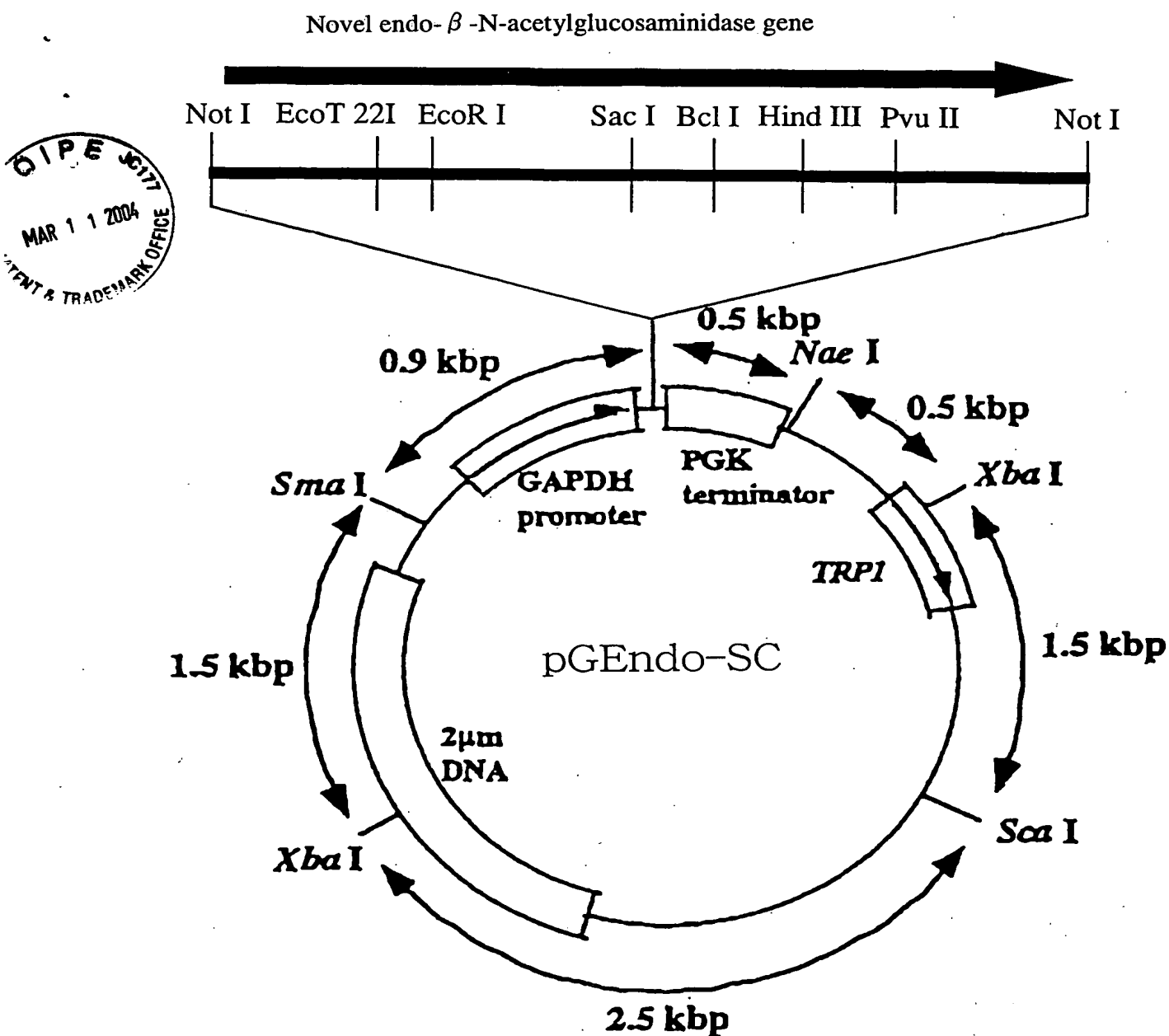
# FIG. 7



1467	1476	1485	1494	1503	1512
AAA GGA TGC TCA TTG CGT TAT ATT TAT AGA ACT TTG TTG ATG AAA GAT GTA AAG					
K G C S L R Y I Y R T L L M K D V K					
1521	1530	1539	1548	1557	1566
TTG ACA GTA GCA TGT CAC TTT TCG TTA AAA ACA AAC GAC TCA GTT AAT TTC TTC					
L T V A C H F S L K T N D S V N F F					
1575	1584	1593	1602	1611	1620
AAG GTA TGG CAG CCA GAT GAA AAT TTC TCT TTT GAA TAT GAT GAT GGA ATG AGA					
K V W Q P D E N F S F E Y D D G M R					
1629	1638	1647	1656	1665	1674
GCC ACT GTT ACA ACT GAA AAT TCT ACC GAA AGC AGA TGC TTT TTA TTA CGT ACA					
A T V T T E N S T E S R C F L L R T					
1683	1692	1701	1710	1719	1728
ACA GAA GAA GAT ACA GGA GAA AAT GAT TGG ATA ACA AAA ACT ATT AAT GTG CCT					
T E E D T G E N D W I T K T I N V P					
1737	1746	1755	1764	1773	1782
GCT GTT CCA GAA GGA AGT CAA TTA TAC ATT ACA AGA CTT GAA GTG AGC GTA GTA					
A V P E G S Q L Y I T R L E V S V V					
1791	1800	1809	1818	1827	1836
TTA GAT ACA GCT GGT TTA GTA GGT CTT GTT AAT CAA GTT ATT GCT TGC TTG GGA					
L D T A G L V G L V N Q V I A C L G					
1845	1854	1863	1872	1881	1890
TAT ATT AGC ATC ATA CCA ACT ATA AAT TCT GGA ATA AAA ACA GAT TCA TCA CGC					
Y I S I I P T I N S G I K T D S S R					
1899	1908	1917	1926	1935	1944
ATT ATT CAG GAT CTC TTT TGG AAA GAT CAG AAA TAT ACC AAA ATC GGA AAA GAA					
I I Q D L F W K D Q K Y T K I G K E					
1953	1962	1971	1980	1989	1998
AGT TTA GAC GAC ATA GCT CAA GAA GAA GTT CAT AGA TAT TAT GGA ACA TTG AAC					
S L D D I A Q E E V H R Y Y G T L N					
2007	2016	2025	2034	2043	2052
TGG GAA AAC ACA GCA AAT GTA GTA AAC GCT TGG GAG GAA ATA GAT TAC TAC AAC					
W E N T A N V V N A W E E I D Y Y N					
2061	2070	2079	2088	2097	2106
GTT TTT TAC AAA GAA AGT GAC GAC TCT GCA ACT CGC ATC TTT TTA GGA ACA GCA					
V F Y K E S D D S A T R I F L G T A					
2115	2124	2133	2142	2151	2160
TTC TGT AAT CAA TTT CGT GTA TCT GGT TTA GAT ATT ATT TTA TCT AAG CTA CCA					
F C N Q F R V S G L D I I L S K L P					
2169	2178	2187	2196	2205	2214
AAG ATA GTT ATT GAA GCT GTT AAC AAA GAA GGA TAC ATC TCT TCA AGT GGT AGC					
K I V I E A V N K E G Y I S S S G S					
2223	2232				
ATA GAT TTG TCA TTA AAC TAG 3'					
I D L S L N *					

Amino acid sequence deduced from the novel Endo- $\beta$ -N-acetylglucosaminidase gene, and the nucleotide sequence of the DNA encoding this amino acid sequence. (Continued)

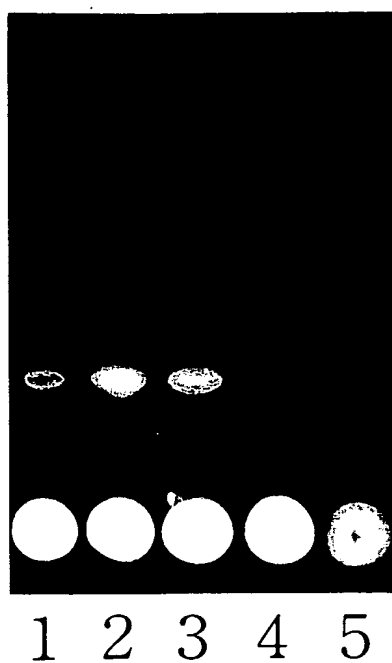
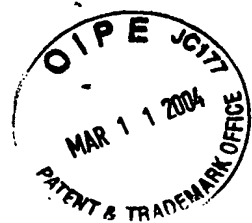
FIG. 8



Structure of expression vector pGEndo-SC for the use in *Saccharomyces cerevisiae*, which comprises a novel endo- $\beta$ -N-acetylglucosaminidase gene.



FIG. 9



Expression of endo- $\beta$ -N-acetylglucosaminidase enzyme in yeast into which an endo- $\beta$ -N-acetylglucosaminidase gene has been introduced.

Lanes 1-3: Cellular extract of *S. cerevisiae* YPH500 (pep4) into which an endo- $\beta$ -N-acetylglucosaminidase gene has been introduced.

Lane 4: Purified endo- $\beta$ -N-acetylglucosaminidase derived from *M. hiemalis*

Lane 5: Cellular extract of *S. cerevisiae* YPH500 (pep4)